HSR Health Services Research

DOI: 10.1111/j.1475-6773.2010.01149.x
INTRODUCTION

Health Services Research in 2020: Data and Methods Needs for the Future

Patricia Pittman

BACKGROUND

Since its inception as a multidisciplinary field of inquiry in the United States nearly 50 years ago, *Health Services Research* (HSR) has experienced major growth spurts during periods of heightened national interest in health care reform (Colby 2008).[1] The massive investments in comparative effective research (CER) and in health information technology (HIT) already seen in the 2009 American Recovery and Reinvestment Act (ARRA) were the first signs that the demand for HSR was likely to greatly expand in coming years. The health reform bill that recently passed confirms this trend, with a significant role for HSR in the design and conduct of pilot evaluations.

In an effort to understand the challenges that the field will face as it enters this historic period, AcademyHealth convened two summits on the future of *HSR*. The first Summit was held in 2007 and focused on the fields' strengths and limitations with regard to its own emerging workforce. Findings were published in the December 2009 Special Section of *HSR* and included documentation of the massive growth in the field's size in the last decade and a series of recommendations regarding the relative roles of researchers trained in HSR and those trained in single disciplines, as well as the importance of linking the HSR training curriculum to a range of new types of HSR employers (Pittman and Holve 2009).

The second Summit, held in June 2009 and the subject of this Special Issue, focused on the data infrastructure that will be needed as the next decade unfolds in order to produce robust and relevant HSR.[2,3,4] The papers commissioned for the Summit examined the changing health policy landscape and its likely effects on the data and methods need of the field, as well as strategies different sectors can adopt to improve the access to and quality of data.

A subset of these papers is included in this Special Issue. They include four papers that focus on data stewardship challenges, spanning data governance (Rosenbaum), privacy and confidentiality (Lane and Schur), incentives to improve data quality (Luft), and data linkages (Bradley, Penberthy, and Devers). Five additional papers examine the data and measurement needs relating to specific policy topics, including patient safety (Goeschel and Provonost), disparities (Bilheimer, Keepel, and Klein), multiple chronic illnesses (Iezzoni), value-based purchasing (McHugh and Joshi), and modeling policy options (Ringel and colleagues).

THE CHANGING DATA ENVIRONMENT

A major focus of the Summit discussion concerned the changing nature of health data and its effects on the way HSR is conducted. The central idea was that, while data have in the past been almost exclusively generated by researchers themselves through audits, surveys, and other collection methods, they are now increasingly produced in spheres that are beyond the control of researchers. Some of these data are generated as a by-product of care delivery and reside in (and are controlled by) health care organizations and health plans. They are produced through increasingly sophisticated technologies, and they also include new sources such as social networking sites that operate in cyberspace and are oriented toward patients' and doctors' use, rather than researchers'. Websites like patientlikeme.com and 23andme.com, and physician sites, such as sermo.com present a new world of opportunity, as well as myriad methodological and governance challenges that the field is only just beginning to contemplate. Most academic researchers have limited experience with data produced through these more novel means.

How researchers will interact with these "naturally occurring data" and engage, or not, with those who control them will have enormous impact on the ways in which the field of HSR develops. Indeed, Summit discussions indicate that some of these changes may be uncomfortable for the field. As Karen Davis remarked in her summary of the discussion, "health services research will need to give up some control." Paul Wallace, of Kaiser Permanente, went so far as to suggest that the changes will bring about a paradigmatic shift in the field—a shift

---

Address correspondence to Patricia Pittman, Ph.D., Associate Professor, Department of Health Policy, George Washington University School of Public Health and Health Services, 2121 K Street NW, Suite 210, Washington, D.C. 20037; e-mail: Patricia.pittman@gwumc.edu

he likens to the transition from an era of food scarcity, with a hunter-gatherer mentality, to an era of abundance associated with agricultural techniques.

Given that these data have not been validated or tested in a research context, a first question is how useful are they for research purposes? Luft suggests that for those data relating to provider performance, incentives can be created to continually improve data quality. For data generated outside health services, such as patient-driven websites, Summit participants urged the field to engage with these initiatives to explore potential partnerships that would help improve data quality and provide guidance on its appropriate use on research.

The changing data landscape gives rise to a second question: how will the high volume of low-cost data impact the traditional research enterprise? In a future scenario that includes regulatory requirements to make data available at no, or low, cost (Rosenbaum), and financial rewards that incentivize continuous improvement of data quality (Luft), the cost of producing research could be dramatically reduced and cost-drivers associated with conducting research could shift. Such a reduction in the costs of research might mean health care organizations become more interested in creating their own HSR units. Similarly, it could mean that the pressure to increase research productivity within university settings could rise, even as researchers struggle to maintain the same levels of funding.

If data quality improves over time and if the cost of producing research is reduced, a third question arises relating to the ability to conduct real-time research that responds to clinical and management level decision-making processes. Even as longer term research, based on static data, will continue to be important for some researchers, it is likely that a subgroup of health services researchers will develop new data analysis methods that are oriented to prospective design using electronic data. To do so, they will need to form partnerships with health care organizations that stand to benefit from the analyses.

The availability of real time data, and data from a broader array of sources, also means that new actors, including clinicians and patients, will have an increased capacity to conduct research. Learning to engage these groups as part of an expanding field of health services research could also challenge academic researchers to broaden the notion of what counts as "research." Over time, the traditional divide between, for example, quality improvement and the health services research, may begin to blur, constituting a continuum of research, rather than two distinct communities of professionals.

Acknowledging the transformations underway for the field of HSR, and taking into account the myriad data and measurement needs put forth in the commissioned papers, Summit participants were asked to reflect on the new data and research policies that will be needed at various levels. The discussion and recommendations that emerged revolved around five themes: access to quality data generated as a by-product of care, increased access and relevance of federal data sets, definitional and measurement challenge, and modernizing and coordinating research policy.

INCREASING ACCESS TO AND QUALITY OF HEALTH SYSTEMS DATA

Participants strongly affirmed the need for a quasi governmental or a public entity, or series of entities, that would be charged with data stewardship. Key to the mission of such an entity would be to determine what data should be made public and under what conditions (e.g., for investigator initiated, government evaluation, or corporate marketing research).

In her paper, Rosenbaum argues that, over time, health information will increasingly be seen as a public good, and that now is the time to establish a data stewardship function, which she suggested includes "a collection of data management methods covering acquisitions, storage, aggregation, and de-identification, and procedures for data release and use" (Rosenbaum). To the extent that health systems data can lead to knowledge about such critical key pubic public policy imperatives as reducing waste and improving care delivery outcomes, and to the extent that the public investment in health care delivery is likely to grow, Rosenbaum asserts that it is politically untenable that the private sector would be able to retain exclusive ownership and create a business of selling data selectively. She argues that, from a legal perspective, for private health care organizations to charge for data access, in effect, makes taxpayers pay twice: once for the care delivery (when patient care is funded by public sources) and again for the documentation of that care.

Acknowledging that there are personal, professional, and business interests at stake, the Summit participants suggested that stakeholder committees be established to provide oversight and guidance to the data stewardship entity, much like the committees that were established to advise the Office of the National Coordinator for Health Information Technology. Provisions relating to HIT and CER in the ARRA are, in fact, constituting the first step in the process of establishing an effective data stewardship function.

A complementary set of recommendations focused on developing market incentives that encourage open access to health care data (Luft).

Recognizing the legitimate fears, including liability risks, that health care organizations have about sharing their own data, and building on a distributed data model, Luft suggests that trustworthy "data consolidators" would ensure the confidentiality of below average performers, while transmitting offers to the best performing practices on behalf of health plans. This would encourage, not only learning from high-quality providers, but a continuous incentive to improve data quality.

In a closely related discussion, Lane and Schur point out the very real danger of re-identifying data on the Internet, as well as linked claims and survey data that can be geographically identified. They affirm that other fields have identified a band of socially tolerable risk in which a combination of four types of protection can be used: technological, statistical, operational, and legal. Consistent with Rosenbaum and Luft, the authors propose that data access should be provided using remote access data enclaves, that is, data consolidators, which also specialize in developing technical knowledge about how best to protect privacy. They also discuss the use of synthetic data as an important strategy for mitigating threats to privacy.

Recommendations to improve access to and quality of health systems data

1. Establish a data stewardship entity that includes stakeholder advisory committees to establish requirements for making certain kinds of data public under certain conditions.
2. Create a trusted data consolidator that gathers and provides de-identified data on provider performance to health plans and to the public, while at the same time offering to transmit health plan rewards to high performers.
3. This data consolidator (or multiple data enclaves) should be charged with taking the technological, statistical, operational, and legal steps necessary to ensure privacy.
4. Public research funders should purchase data remaining in the private sector that are needed for publicly funded research only once, and then make it available for multiple projects.

A final point on the issue of data access was that currently research funders pay for private data separately for each study. In the future, any data that remain in private sector hands should be purchased by governmental research agencies just once with an understanding that it may be used for multiple studies. An advantage of this model is that it could also help reduce the variation in interpretations among institutional review boards, which continue to frustrate health services researchers, especially those attempting to conduct multisite research.

Summit participants discussed the need for clearer guidance on existing data privacy regulations, as well as the need for new legislation that would shift the burden of misuse of data from providers of the data to the exclusive responsibility of researchers, with corresponding penalties.

IMPROVING FEDERAL DATA SETS

Summit papers and discussions also focused on the need to improve access to and quality of government survey data. Bradley, Penberthy, and Deverspoint point out that researchers sometimes face lengthy bureaucratic processes to obtain permission to use data sets, or may only be allowed access to data if they conduct their analyses, while physically located at the agency's offices. Summit participants called for the creation of a single online portal for federal health data sets, as has already occurred in other areas of scientific inquiry. There was also a strong interest in streamlining the data use approval processes.

Bradley, Keepel, and Klein also document the challenges of data linkages and urge standardization of data definitions among the multiple federal data sets to ease the process and to boost the ability to link datasets. They argue that too few researchers are trained to link data and recommend additional training and the creation of a mechanism to allow linked data to be shared. Participants suggested that this "mechanism" could be some form of a "national library" that would operate in conjunction with the data consolidator or remote data enclaves. It would maintain linked data from federal surveys and/or private data sets in a protected space. This would reduce the time and costs associated with subsequent linking efforts, thereby diminishing research costs.

Several of the policy papers stressed that, as the health system changes, there is a continuous need to develop new units of analysis in federal surveys. Salient examples include:

- data on teams of health care practice, rather than data on single professions,
- data on organizations other than hospitals, and
- data on health plans.

Participants interested in tracking the uninsured suggested that surveys such as the Health and Retirement Survey should include a full range of health insurance options, including prices, percent of family budgets spent on health insurance and other goods, reasons for un-insurance, self-assessment of risks, valuation of insurance, mental health and financial wealth dynamics, and productivity measures.

Others suggested that differences in labor and insurance markets, as well as provider characteristics, make it important to include variables that allow geographic variation to be monitored.

Bilheimer, Keepel, and Klein describe the limitations of many of the current surveys and administrative data systems in tracking differences among population subgroups. They outline several strategies to address those problems, including larger surveys, more targeted surveys, and standards for patient demographic data in HIT systems. They also suggest that greater use of linked survey and administrative data could facilitate the analysis of the social, economic, and environmental determinants of health.

Several papers call for more survey panel data that follow individuals over time and across sites of care. Girosi and his colleagues describe the challenges in the area of modeling policy alternatives. They propose that a health-focused, nationally representative, longitudinal study on individuals be fielded to clarify behavioral responses to policy changes.

Recommendations to improve access to and quality of federal survey data

1. A single online portal for all federal surveys should be created.
2. To facilitate data linkages, the federal government should pursue opportunities to standardize surveys, provide more training opportunities for researchers, and generate a new mechanism to allow data linked for one research project to remain linked for future use, for example, a national library.
3. The federal government should consider adding new units of analysis to existing surveys, for example, a national provider ID, data on teams of practice, on nonphysician clinicians, on health care organizations that are not hospitals, on health plans, and on special population groups, in particular those with multiple chronic diseases and vulnerable populations, including children, and developing new surveys that follow individuals over time and across sites of care are needed.

RESEARCH POLICY

Beyond the data governance issues discussed in the previous two sections, there are also implications of this changing landscape for those in a position to help define research priorities, including HSR funders, journals, and professional associations such as AcademyHealth. Themes that emerged from the papers and the Summit deliberations point to the importance of investing in the development of new definitions and measurements in emerging research areas, greater support for implementation research, and incentives to form partnerships between academic researchers and delivery systems to design and analyze data that are generated as a by-product of care.

Iezzoni's paper brings attention to the measurement challenges associated with the growing prevalence of multiple chronic conditions and disabilities. She describes nascent developments in this area, including the advantages of adopting the ICD-10-CM in 2013 and the contributions of the International Classification of Functioning, Disability and Health. These instruments are key to being able to develop performance measures that take into account this large population, as well as the ability to study the effects of multiple chronic conditions and disabilities on therapeutic effectiveness and the development of appropriate treatment options for this special population. She goes on to suggest, however, that significant incentives and training will be required before these instruments can be programmed into electronic medical records and leveraged for research.

Bilheimer, Keepel, and Klein also focus on the need for better measures and definitions in the area of disparities research. They propose a series of definitions relating to equity, equality, disparities, and burden of disease that they believe could help to clarify measurement challenges. Similarly, Goeschel and Provonost make the case for investing in measurement development in the area of patient safety research.

In their paper on value-based purchasing, McHugh and Joshi explore the critical role of implementation research, that is, research that assesses the process of applying research-based knowledge to policy and practice stress. While academic researchers bear some responsibility for the sparse attention given to this type of research, Summit participants agreed that federal agencies that support HSR, as well as the journals that publish HSR, could play a more active leadership role in expanding the concept of HSR to actively embrace these types of evaluations.

Finally, there was an interest among participants in exploring ways that funders could alter current academic incentives, such that researchers would be rewarded for partnerships with health systems. Similarly, incentives could be created on the provider side to encourage them to seek out researchers. Such incentives might include funding mechanisms for quick turnaround studies that have immediate application to care delivery improvements, or other more innovative types of rewards such as seed funding to begin a line of collaborative arrangements that encourage exchange among researchers and practitioners on research needs as well as results. Research funding could be tied to this user input mechanism. In addition, journals, which tend to favor papers with generalizable findings and to be less interested in implementation research and methods development, could be encouraged through grants to expand their traditional preferences.

Recommendations for researchers, funders, and journals

1. In areas of HSR that are relatively new and growing in importance, funders and researchers should focus on investing in the development of standard definitions and measurements that can form the foundation for future inquiry.
2. Researchers need to embrace new partnerships with providers that may involve having less control over research questions and data than has been the case in the past.
3. Research funders and journals should consider expanding definitions of research to support implementation research and open collaborative environments that facilitate participation of research users in the design, conduct, and application of research.

CONCLUSIONS

Health reform has provided an historic opportunity to demonstrate the important role HSR can play informing health policy and practice. It is incumbent upon those both inside and outside of government who want the field to succeed to consider ways to strengthen the research infrastructure. Perhaps more than any other infrastructure component, data define the possibilities and the limits of HSR.

The papers included in this special issue and the recommendations derived from the Summit set forth a range of strategies to strengthen the data infrastructure that will be needed to produce HSR in the future These recommendations are relevant to several groups of actors, including government agencies, public and private providers, educational institutions, public and private research funders, and of course health services researchers themselves along with their professional association, AcademyHealth. While each of the three areas of recommendations will require further elaboration, we hope that together they provide a roadmap for discussions among these various sectors on ways to strengthen the data infrastructure needed for HSR.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This initiative was led by AcademyHealth and supported by grants from the Robert Wood Johnson Foundation and the Commonwealth Fund. Guidance was provided by the project's Advisory Committee[5] and AcademyHealth staff: Sharon Arnold, Erin Holve, Alison Rein, and Michael Gluck.

NOTES

1. Health services research has been defined as a "multidisciplinary field of scientific investigation that studies how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and ultimately our health and well-being" (Lohr and Steinwachs 2002).
2. See http://www.academyhealth.org/About/content.cfm?ItemNumber=2529
3. Participants in the Summit: Sharon Arnold, Laurence Baker, Anne Beal, Mark Benton, Aman Bhandari, Linda Bilheimer, Lynn Blewett, Cathy Bradley, Martin Brown, Martin Brown, Peter Buerhaus, Francis Chesley, David Colby, Sarah Dash, Karen Davis, Kelly Devers, Carol Diamond, Mark Doescher, Bryan Dowd, Christine Eibner, Seth Eisen, Erin Fraher, Richard Glicklich, Michael Gluck, Christine Goeschel, Marsha Gold, Jonathan Gruber, David Helms, Erin Holve, Grant Huang, Lisa Iezzoni, Maulik Joshi, KenenthKenneth Keppel, Richard Klein, Mahesh Krishnan, Julia Lane, Teresa Lee, Harold Luft, William Marder, Megan McHugh, Klim McPherson, Penny Mohr, Vincent Mor, Farzad Mostashari, Patricia Pittman, Steven Pizer, Peter Pronovost, Alison Rein, Sara Rosenbaum, Joachim Roski, Claudia Schur, Pauline Sieverding, Edward Sondik, Joanne Spetz, Walter Stewart, Michael Stoto, Paul Wallace, and Claudia Williams.
4. For a full list of papers see http://www.academyhealth.org/Events/content.cfm?ItemNumber=2529
5. Advisory Committee members: David Abrams, American Legacy Foundation; Linda Bilheimer, National Center for Health Statistics; Francis Chesley, Agency for Healthcare Research and Quality; Carolyn Clancy, Agency for Healthcare Research and Quality; David Colby, Robert Wood Johnson Foundation; Kelly Devers, Virginia Commonwealth University; Bryan Dowd, University of Minnesota; Richard Gliklich, Outcome Sciences; Grant Huang, VA Cooperative Studies Program; Susan Law, Canadian Health Services Research Foundation; Patricia Mabry, National Institutes of Health; Vincent Mor, Brown University School of Medicine; Pauline Sieverding, Department of Veterans Affairs; Michael Stoto, Georgetown University School of Nursing & Health Studies; Philip Wang, National Institute of Mental Health.

REFERENCES

Colby, D. 2008. "Health Services Research." In *To Improve Health and Health Care Vol. XI, Robert Wood Johnson Foundation Anthology*, edited by Isaacs, S., and D. Colby, p. 2. San Francisco, CA: Jossey-Bass.

Lohr, K., and D. Steinwachs. 2002. "Health Services Research: An Evolving Definition of the Field." *Health Services Research* 37 (1): 7–9.

Pittman, P., and E. Holve. 2009. "The Health Services Researcher of 2020: A Summit to Assess the Fields' Workforce Needs." *Health Services Research* 44 (6): 2198–213.